Electronic Transport in DNA

Daphne Klotsa, Rudolf A. Römer, Matthew S. Turner
Physics Department and Centre for Scientific Computing,
University of Warwick, Coventry CV4 7AL, U.K.
(Dated: *Revision* : 1.45, compiled February 9, 2008)

We study the electronic properties of DNA by way of a tight-binding model applied to four particular DNA sequences. The charge transfer properties are presented in terms of localisation lengths, crudely speaking the length over which electrons travel. Various types of disorder, including random potentials, are employed to account for different real environments. We have performed calculations on poly(dG)-poly(dC), telomeric-DNA, random-ATGC DNA and λ-DNA. We find that random and λ-DNA have localisation lengths allowing for electron motion among a few dozen base pairs only. A novel enhancement of localisation lengths is observed at particular energies for an increasing binary backbone disorder. We comment on the possible biological relevance of sequence dependent charge transfer in DNA.

PACS numbers: 72.15.Rn, 87.15.Cc, 73.63.-b

I. INTRODUCTION

The question of whether DNA conducts electric charges is intriguing to physicists and biologists alike. The suggestion that electron transfer/transport in DNA might be biologically important has triggered a series of experimental and theoretical investigations [5, 17, 20, 31, 35, 54]. Processes that possibly use electron transfer include the function of DNA damage response enzymes, transcription factors or polymerase co-factors all of which play important roles in the cell [2]. Indeed there is direct evidence [9] that MutY — a DNA base excision repair enzyme with an $[4Fe4S]^+$ cluster of undetermined function — takes part in some kind of electron transfer as part of the DNA repair process [36, 46]. This seems consistent with studies in which an electric current is passed through DNA revealing that damaged regions have significantly different electronic behaviour than healthy ones [9].

For physicists, the continuing progress of nanotechnologies and the consequent need for further size miniaturisation makes the DNA molecule an excellent candidate for molecular electronics [6, 13, 23, 45]. DNA might serve as a wire, transistor, switch or rectifier depending on its electronic properties [16, 20, 44].

In its natural environment, DNA is always in liquid solution and therefore experimentally one can study the molecule either in solution or in artificially imposed dry environments. In solution experiments DNA can be chemically processed to host a donor and an acceptor molecule at different sites along its long axis. Photo-induced charge transfer rates can then be measured whilst the donor/acceptor molecules, the distance and the sequence of DNA that lies between them are varied. The reactions are observed to depend on the type of DNA used, the intercalation, the integrity of the intervening base pair stack and, albeit weakly, on the molecular distance [5, 9, 17, 35, 52].

Direct conductivity measurements on dry DNA have also been preformed in the past few years. The remarkable diversity that characterises the results seems to arise from the fact that many factors need to be experimentally controlled. These include methods for DNA alignment and drying, the nature of the devices used to measure the conductivity, the type of metallic contacts and the sequence and length of the DNA. DNA has been reported to be an insulator [10], an ohmic conductor [3, 21, 32, 34, 45] and a semiconductor [43]. Theoretically, single-step super exchange [31] and multi-step hopping [8] models have provided interpretations of solution experiments. For experiments in dry DNA, several additional approaches such as variable range hopping [57], one-dimensional quantum mechanical tight-binding models [13, 47, 48, 55, 58, 59] and non-linear methods [12, 39] have also been proposed.

Despite the lack of a consistent picture for the electronic properties of DNA, one conclusion has been established: the environment of the DNA impacts upon its structural, chemical and thus probably also electronic properties. Both theoretical and experimental studies show that the temperature and the type of solution surrounding DNA have a significant effect on its structure and shape [4, 11, 57]. The effect of the environment is a key one to this report, where the environmental fluctuations are explicitly modelled as providing different types of disorder.

In this work, we focus on whether DNA, when treated as a quantum wire in the fully coherent low-temperature regime, is conducting or not. To this end, we study and generalise a tight-binding model of DNA which has been shown to reproduce experimental [13] as well as *ab-initio* results [15]. A main feature of the model is the presence of sites which represent the sugar-phosphate backbone of DNA but along which no electron transport is permissible. We measure the "strength" of the electronic transport by the *localisation length* ξ, which roughly speaking parametrises whether an electron is confined to a certain region ξ of the DNA (insulating behaviour) or can proceed across the full length L ($\leq \xi$) of the DNA molecule (metallic behaviour).

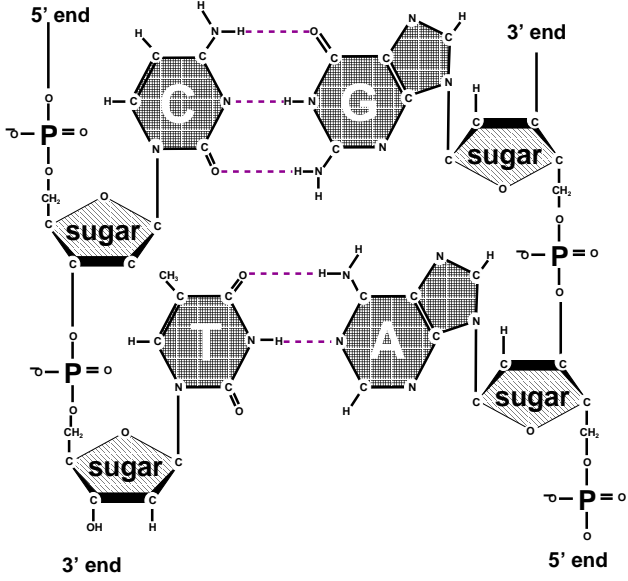Sections II–III introduce our models and the numerical

FIG. 1: The chemical composition of DNA with the four bases Adenine, Thymine, Cytosine, Guanine and the backbone. The backbone is made of phosphorylated sugars shown in yellow and brown.

approach. In section V, we show that DNA sequences with different arrangement of nucleotide bases Adenine (A), Cytosine (C), Guanine (G) and Thymine (T) exhibit different ξ's when measured, e.g. as function of the Fermi energy E. The influence of external disorder, modelling variants in the solution, bending of the DNA molecule, finite-temperature effects, etc., is studied in section VI where we show that, surprisingly, the models support an increase of ξ when disorder is increased. We explain that this effect is linked to the existence of the backbone sites.

II. TIGHT-BINDING MODELS FOR DNA WITH A GAP IN THE SPECTRUM

A. The Fishbone model

DNA is a macro-molecule consisting of repeated stacks of bases formed by either AT (TA) or GC (CG) pairs coupled via hydrogen bonds and held in the double-helix structure by a sugar-phosphate backbone. In Fig. 1, we show a schematic drawing. In most models of electronic transport [13, 60] it has been assumed that the transmission channels are along the long axis of the DNA molecule [61] and that the conduction path is due to π-orbital overlap between consecutive bases [52]; density-functional calculations [37] have shown that the bases, especially Guanine, are rich in π-orbitals. Quantum mechanical approaches to the problem mostly use strictly one-dimensional (1D) tight-binding models [47, 48, 55, 58, 59].

Of particular interest to us is a quasi-1D model [13] which includes the backbone structure of DNA explicitly
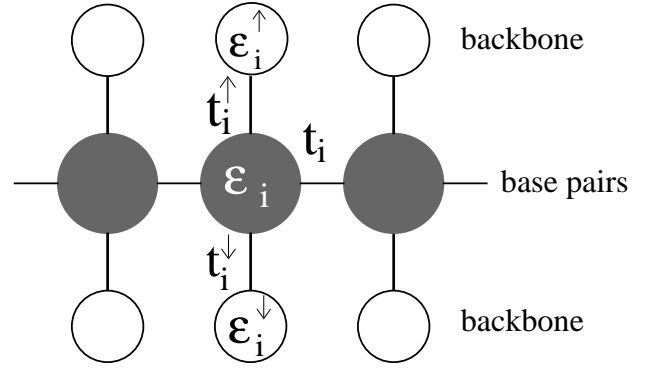

FIG. 2: The fishbone model for electronic transport along DNA corresponding to the Hamiltonian given in Eq. (1). Lines denote hopping amplitudes and circles give the central (grey) and backbone (open) sites.

and exhibits a semiconducting gap. This *fishbone model*, shown in Fig. 2, has one central conduction channel in which individual sites represent a base-pair; these are interconnected and further linked to upper and lower sites, representing the backbone, but are *not* interconnected along the backbone. Every link between sites implies the presence of a hopping amplitude. The Hamiltonian for the fishbone model (H_F) is given by:

$$H_F = \sum_{i=1}^{L} \sum_{q=\uparrow,\downarrow} (-t_i|i\rangle\langle i+1| - t_i^q|i,q\rangle\langle i| + \varepsilon_i|i\rangle\langle i| + \varepsilon_i^q|i,q\rangle\langle i,q|) + h.c. \quad (1)$$

where t_i is the hopping between nearest-neighbour sites $i, i+1$ along the central branch, t_i^q with $q = \uparrow, \downarrow$ gives the hopping from each site on the central branch to the upper and lower backbone respectively. Additionally, we denote the onsite energy at each site along the central branch by ε_i and the onsite energy at the sites of the upper and lower backbone is given by ε_i^q, with $q = \uparrow\downarrow$. L is the number of sites/bases in the sequence. The model (1) clearly represents a dramatic simplification of DNA. Nevertheless, in Ref. [13] it had been shown that this model when applied to an artificial sequence of repeated GC base pairs, poly(dG)-poly(dC) DNA, reproduces experimental data current-voltage measurements when $t_i = 0.37eV$ and $t_i^q = 0.74eV$ are being used. Therefore, we will assume $t_i^q = 2t_i$ and set the energy scale by $t_i \equiv 1$ for hopping between GC pairs. In what follows we will adopt energy units in which $eV = 1$ throughout.

For natural DNA sequences, we need to know how the hopping amplitudes vary as the electron moves between like pairs, i.e. from GC to GC or from AT to AT, and unlike pairs, i.e., from GC to AT and vice versa. We choose $t_i = 1$ between identical and matching bases (e.g. AT/TA, GC/CG). Assuming that the wavefunction overlap between consecutive bases along the DNA strand is weaker between unlike and non-matching bases (AT/GC, TA/GC, etc.) we thus choose $1/2$.
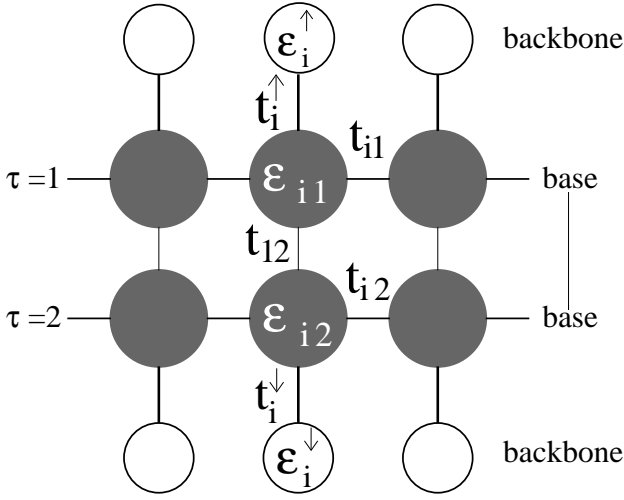
FIG. 3: The ladder model for electronic transport along DNA. The model corresponds to the Hamiltonian (2).

B. The Ladder model

We performed semi-empirical calculations on DNA base pairs and stacks using the SPARTAN quantum chemistry software package [1]. The results have shown that the relevant electronic states of DNA (highest-occupied and lowest-unoccupied molecular orbitals with and without an additional electron) are localised on one of the bases of a pair only. The reduction of the DNA base-pair architecture into a single site per pair, as in the fishbone model (1), is obviously a highly simplified approach. As an improvement on this we model each base as a distinct site where the base pair is then weakly coupled by the hydrogen bonds. The resulting 2-channel model is shown in Fig. 3. This *ladder* model is a planar projection of the structure of the DNA with its double-helix unwound. We note that results for electron transfer also suggest that the transfer proceeds preferentially down one strand [25]. There are two central branches, linked with one another, with interconnected sites where each represents a complete base and which are additionally linked to the upper and lower backbone sites. The backbone sites as in the fishbone model are not interconnected. The Hamiltonian for the ladder model is given by

$$
\begin{aligned}
H_L = \sum_{i=1}^{L} \Bigg[& \sum_{\tau=1,2} (t_{i,\tau}|i,\tau\rangle\langle i+1,\tau| + \varepsilon_{i,\tau}|i,\tau\rangle\langle i,\tau|) \\
& + \sum_{q=\uparrow,\downarrow} (t_i^q|i,\tau\rangle\langle i,q(\tau)| + \varepsilon_i^q|i,q\rangle\langle i,q|) \\
& + t_{1,2}|i,1\rangle\langle i,2| \Bigg] + h.c.
\end{aligned} \tag{2}
$$

where $t_{i,\tau}$ is the hopping amplitude between sites along each branch $\tau = 1, 2$ and $\varepsilon_{i,\tau}$ is the corresponding onsite potential energy. t_i^q and and ε_i^q as before give hopping amplitudes and onsite energies at the backbone sites.

Also, $q(\tau) = \uparrow, \downarrow$ for $\tau = 1, 2$, respectively. The new parameter t_{12} represents the hopping between the two central branches, i.e., perpendicular to the direction of conduction. SPARTAN results suggest that this value, dominated by the wave function overlap across the hydrogen bonds, is weak and so we choose $t_{12} = 1/10$.

C. Including disorder

In order to study the transport properties of DNA, we could now either use artificial DNA (poly(dG)-poly(dC) [43], random sequences of A,T,G,C [38, 56], etc.) or natural DNA (bacteriophage λ-DNA [37], etc.). The biological content of the sequence would then simply be encoded in a specific sequence of hopping amplitudes 1 and 1/2 between like and unlike base-pair sequences. However, in vivo and most experimental situations, DNA is exposed to diverse environments and its properties, particularly those related to its conformation, can change drastically depending on the specific choice. The solution, thermal effects, presence of binding and packaging proteins and the available space are factors that alter the structure and therefore the properties that one is measuring [4, 57]. Clearly, such dramatic changes should also be reflected in the electronic transport characteristics. Since it is precisely the backbone that will be most susceptible to such influences, we model such environmental fluctuations by including variations in the onsite potentials $\varepsilon_{i,q}$.

Different experimental situations will result in a different modification of the backbone electronic structure, and we model this by choosing different distribution functions for the onsite potentials, ranging from uniform disorder $\varepsilon_{i,q} \in [-W/2, W/2]$, to Gaussian disorder and on to binary disorder $\varepsilon_{i,q} = \pm W/2$. W is a measure for the strength of the disorder in all cases. Particularly the binary disorder model can be justified by the localisation of ions or other solutes at random positions along the DNA strand [4].

D. Effective models and the energy gap

Due to the non-connectedness of the backbone sites along the DNA strands, the models (1) and (2) can be further simplified to yield models in which the backbone sites are incorporated into the electronic structure of the DNA. The effective fishbone model is then given by

$$
\begin{aligned}
\tilde{H}_F = \sum_{i=1}^{L} & -t_i|i\rangle\langle i+1| + h.c. \\
& + \left[\varepsilon_i - \sum_{q=\uparrow,\downarrow} \frac{(t_i^q)^2}{\varepsilon_i^q - E} \right] |i\rangle\langle i| .
\end{aligned} \tag{3}
$$

Similarly, the effective ladder model reads as

$$
\begin{aligned}
\tilde{H}_L =\ & \sum_{i=1}^{L} t_{1,2}|i,1\rangle\langle i,2| + \sum_{\tau=1,2} t_{i,\tau}|i,\tau\rangle\langle i+1,\tau| \\
& + \left[\varepsilon_{i,\tau} - \frac{\left(t_i^{q(\tau)}\right)^2}{\varepsilon_i^{q(\tau)} - E} \right]|i,\tau\rangle\langle i,\tau| \\
& + h.c. \quad .
\end{aligned}
\tag{4}
$$

In these two models, the backbone has been incorporated into an energy-dependent onsite potential on the main DNA sites. This re-emphasises that the presence of the backbone influences the local electronic structure on the DNA bases and similarly, any variation in the backbone disorder potentials $\varepsilon_i^{\uparrow,\downarrow}$ will results in a variation of *effective* onsite potentials as given in the brackets of Eqs. (3) and (4).

Both models allow to quickly calculate the gap of the completely ordered system (all onsite potentials zero) by assuming that the lowest-energy state $\psi = \sum_i \psi_{i,(\tau)}|i,(\tau)\rangle$ in each band corresponds to constant ψ_i ($\psi_{i,\tau}$) whereas for the highest-energy states, a checkerboard pattern is obtained with $\psi_i = \psi_{i+1}$ ($\psi_{i,\tau} = -\psi_{i+1,\tau}$, $\psi_{i,1} = -\psi_{i,2}$). For the fishbone model, this shows that, e.g. $E_{\min,\mp} = -t_i \mp \sqrt{t_i^2 + t_{i,\uparrow}^2 + t_{i,\downarrow}^2}$ and $E_{\max,\mp} = t_i \mp \sqrt{t_i^2 + t_{i,\uparrow}^2 + t_{i,\downarrow}^2}$. For the chosen set of hopping parameters for (3) and (4), this gives $E_{\min,\mp} = -4,2$ and $E_{\max,\mp} = -2,4$ for the fishbone model and $E_{\min,\mp} \approx -3.31, 1.21$ and $E_{\max,\mp} = -1.21, 3.31$ for the ladder model.

III. THE NUMERICAL APPROACH AND LOCALISATION

There are several approaches suitable for studying the transport properties of the models (1) and (2) and these can be found in the literature on transport in solid state devices, or, perhaps more appropriately, quantum wires. Since the variation in the sequence of base pairs precludes a general solution, we will use two methods well-known from the theory of disordered systems [50].

The first method is the iterative transfer-matrix method (TMM) [26, 29, 30, 40, 41] which allows us in principle to determine the localisation length ξ of electronic states in systems with cross sections $M = 1$ (fishbone) and 2 (ladder) and length $L \gg M$, where typically a few million sites are needed for L to achieve reasonable accuracy for ξ. However, in the present situation we are interested in finding ξ also for viral DNA strands of typically only a few ten thousand base-pair long sequences. Thus in order to restore the required precision, we have modified the conventional TMM and now perform the TMM on a system of fixed length L_0. This modification has been previously used [22, 33, 49] and

may be summarised as follows: After the usual forward calculation with a global transfer matrix $\mathcal{T}_{L_0}$, we add a backward calculation with transfer matrix $\mathcal{T}_{L_0}^{\mathrm{b}}$. This forward-backward-multiplication procedure is repeated K times. The effective total number of TMM multiplications is $L = 2KL_0$ and the global transfer-matrix is $\tau_L = \left(\mathcal{T}_{L_0}^{\mathrm{b}}\mathcal{T}_{L_0}\right)^K$. It can be diagonalised as for the standard TMM with $K \to \infty$ to give $\tau_L^\dagger \tau_L \to \exp[\mathrm{diag}(4KL_0/\xi_\tau)]$ with $\tau = 1$ or $\tau = 1, 2$ for fishbone and ladder model, respectively. The largest $\xi_\tau \forall \tau$ then corresponds to the localisation lengths of the electron on the DNA strand and will be measured in units of the DNA base-pair spacing (0.34 nm).

The second method that we will use is the recursive Green function approach pioneered by MacKinnon [27, 28]. It can be used to calculate the dc and ac conductivity tensors and the density of states (DOS) of a d-dimensional disordered system and has been adopted to calculate all kinetic linear-transport coefficients such as thermoelectric power, thermal conductivity, Peltier coefficient and Lorentz number [51].

The main advantage of both methods is that they work reliably (i) for short DNA strands ranging from 13 (DFT studies [37]) base pairs up to 30 base pairs length which are being used in the nanoscopic transport measurements [15] as well as (ii) for somewhat longer DNA sequences as modelled in the electron transfer results and (iii) even for complete DNA sequences which contain, e.g. for human chromosomes up to 245 million base pairs [2].

IV. DNA SEQUENCES

The exact arrangement of the four bases A, T, G, C determines the nature and function of its associated DNA strand such as the chemical composition of the proteins which are encoded. While previous studies have aimed to elucidate whether DNA conducts at all, we shall also focus our attention to investigate how different DNA sequences, be they artificial or naturally occurring, "conduct" charge differently. Thus we study a set of different DNA.

A convenient starting point for most electronic transport studies [44] is the aforementioned poly(dG)-poly(dC) sequence, which corresponds to a simple repetition of a GC (or CG) pair. Note that within our models, there is no difference between GC and CG pairs. Although not occurring naturally, such sequences can be synthesised easily. Another convenient choice of artificial DNA strand is a simple *random* sequence of the four bases, which we construct with equal probability for all 4 bases. However, they are not normally used in experiments.

As DNA samples existing in living organisms, we shall use λ-DNA of the bacteriophage virus [Bacteriophage lambda] which has a sequence of 48502 base pairs. It corresponds to a bacterial virus and is biologically very well characterised. We also investigate the

29728 bases of the SARS virus [SARS]. Telomeric DNA is a particular buffer part at the beginning and ends of of DNA strands for eukaryote cells [2]. In mammals it is a Guanine rich sequence in which the pattern TTAGGG is repeated over thousands of bases. Its length is known to vary widely between species and individuals but we assume a length of 6000 base-pairs. Last, we show some studies of centromeric DNA for chromosome 2 of yeast with 813138 base pairs [CEN2]. This DNA is also reportedly rich in G bases and has a high rate of repetitions which should be favourable for electronic transport.

Initially, we will compute transport properties for complete DNA sequences, i.e. including and not differentiating between coding and non-coding sequences (this distinction applies to the naturally occurring DNA strands only). However, we will later also study the difference between those two different parts of a given DNA. We emphasise that while non-coding DNA suffers from the label of "junk", it is now known to play several important roles in the functioning of DNA [2].

Before leaving the description of our DNA sequences, we note that occasionally, we show results for "scrambled" DNA. This is DNA with the same number of A, T, C, G bases, but with their order randomised. Clearly, such sequences contain the same set of electronic potentials and hopping variations, but would perform quite differently if released into the wild. A comparison of their transport properties with those from the original sequence thus allows to measure how important the exact fidelity of a sequence is.

V. RESULTS FOR CLEAN DNA

Let us start by studying the localisation properties of DNA without any onsite disorder either at $\varepsilon_{i,\tau}$ or at $\varepsilon_{i,q}$. For a poly(dG)-poly(dC) sequence, both fishbone and ladder model produce two separate energy bands between the extremal values computed at the end of section II D. Within these energy bands, the electronic states are extended with infinite localisation length ξ as expected. Outside the bands, transport is exponentially damped due to an absence of states and the ξ values are very close the zero. In Fig. 4 the resulting *inverse* localisation lengths are shown. These are zero for the extended states in the two bands, but finite outside, showing the quick decrease of the localisation lengths outside the bands. In Fig. 5, we show the same data but now plot the localisation length itself. We see that the energy gap observed previously [13] for the poly(dG)-poly(dC) sequence in the fishbone model remains. The difference with respect to the ladder model is a slight renormalisation of the gap width. The localisation lengths of poly(dG)-poly(dC) DNA tend to infinity, meaning that the sequence is perfectly conducting. This is expected due to its periodic electronic structure.

Turning our attention to the other three DNA sequences, we find that telomeric DNA also gives rise to
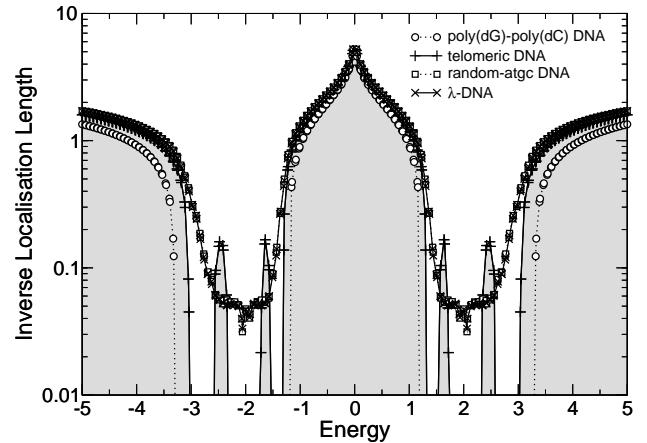

FIG. 4: Plot of the inverse localisation lengths ξ as a function of Fermi energy for the ladder model (4 and four DNA sequences as well as for the fishbone model with a poly(dG)-poly(dC) sequence. The data for telomeric DNA has been shaded for clarity. Lines are guides to the eye only.
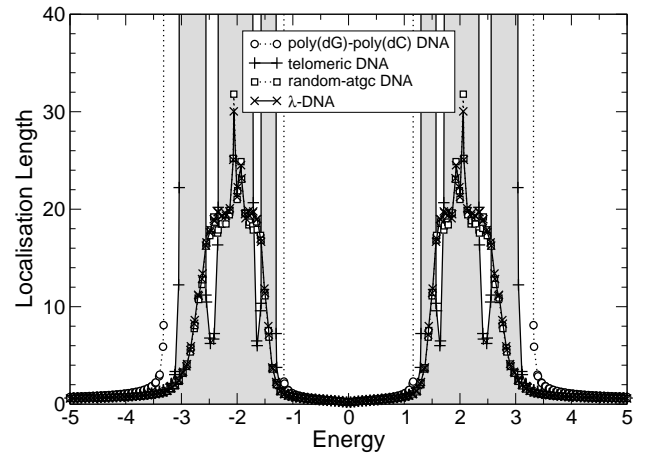

FIG. 5: Localisation lengths as a function of energy for poly(dG)-poly(dC), telomeric, random-ATGC, and λ-DNA as described in the text. The spectrum is symmetric in energy. The data for telomeric DNA has been shaded for clarity. Lines are guides to the eye only.

perfect conductivity like poly(dG)-poly(dC) DNA. But due to its structure of just 6 repeating base pairs, there is a further split of each band into 3 separate sub-bands. They may be calculated like in section II D. We would like to point out that it may therefore be advantageous to use the naturally occurring telomeric parts of DNA sequences as prime, in-vivo candidates when looking for good conductivity in a DNA strand.

The structure of the energy dependence for the random-ATGC and the λ-DNA is very different from the preceding two sequences, but it is quite similar between just these two. The biological content of the DNA sequences is — within the description by our quantum models — just a sequence of binary hopping elements between like and unlike base pairs. Thus the models are

related to the physics of random hopping models [7, 19] and in agreement with these, we see a Dyson peak [18] in the centre of each sub-band. Furthermore, we see that the range of energies for which we observe non-zero localisation lengths is increased into the gap and for large absolute values of the energy. This is similar to the broadening of the single energy band for the Anderson model of localisation [50]. The localisation lengths, which roughly equal the average distance an electron would be able to travel (conduct), are close to the distance of 20 bases within the band, with a maximum of ~ 30 bases at the centre of each band. Note that this result is surprisingly good — given the level of abstraction used in the present models — when compared to the typical distances over which electron transfer processes have been shown to be relevant [9, 17, 25, 31, 35, 52, 54].

VI.   RESULTS FOR DISORDERED DNA

A.   DNA randomly bent or at finite temperatures

As argued before, environmental influences on the transport properties of DNA are likely to influence predominantly the electronic structure of the backbone. Within our models, this can be captured by adding a suitable randomness onto the backbone onsite potentials ε_i^q. In this fashion, we can model for example the influence of a finite-temperature [11] and thus a coupling to phonons [24]. We emphasise however, that in order for our localisation results — which rely on quantum mechanical interference effects — to remain valid, the phase breaking lengths should stay much larger than the sequence lengths. Thus the permissible temperature range is a few K only. The bending of DNA is another possibility which can be modelled by a local, perhaps regular, change in ε_i^q along the strand. Another important aspect is the change in ε_i^q due to the presence of a solution in which DNA is normally immersed.

All these effects can be modelled in a first attempt by choosing an appropriate distribution function $P(\varepsilon_i^q)$. Let us first choose uniform disorder with $\varepsilon_i^q \in [-W/2, W/2]$. In Fig. 6 we show the results for all 4 DNA sequences as a function of energy for $W = 1$. Comparing this to Fig. 5, we see that now all localisation lengths are finite; poly(dG)-poly(dC) and telomeric DNA having localisation lengths of a few hundreds and a few tens of bases, respectively. The localisation lengths for random-ATGC and λ-DNA are only slightly reduced. In all cases, the structure of 2 energy bands remains. Furthermore, $W = 1$ already represents a sizable broadening of about $1/2$ the width of each band. Thus although the localisation lengths are finite compared to the results of section V, they are still larger than the lengths of the DNA strands used in the nano-electric experiments, implying finite conductances. We remark that the Dyson peaks have vanished as expected [19]. We plot the DOS for λ-DNA in Fig. 6 which clearly indicates the 2 bands.
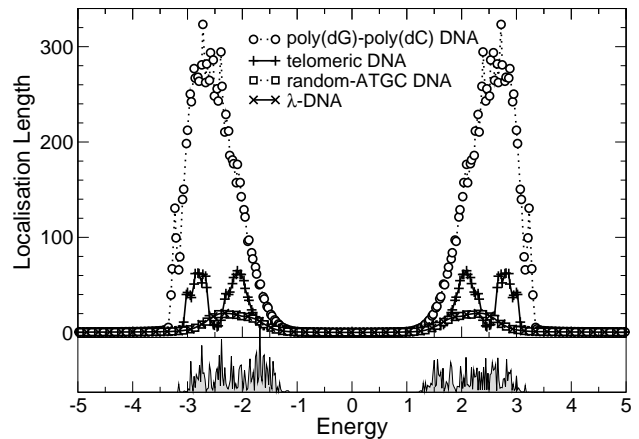

FIG. 6: Top: Energy dependence of the localisation lengths, $\xi(E)$, for poly(dG)-poly(dC), telomeric, random-ATGC and λ-DNA in the presence of *uniform* backbone disorder with $W = 1$. Only every 2nd and 5th symbol is shown for random-ATGC and λ-DNA, respectively. Bottom: DOS for λ-DNA using the same parameters as in the top panel.
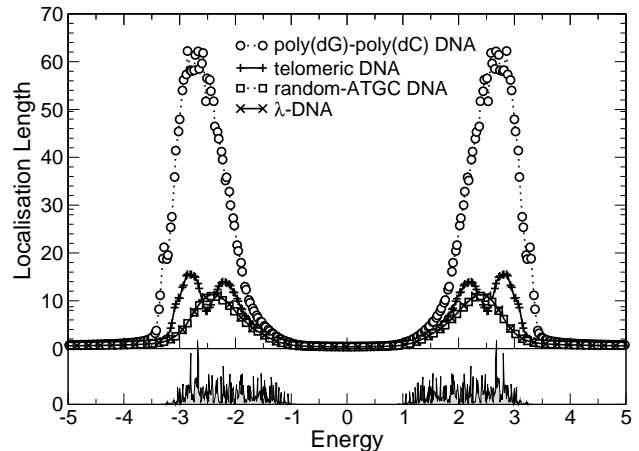

FIG. 7: Top: $\xi(E)$ as in Fig. 6 but with $W = 2$. Only every 2nd and 5th symbol is shown for random-ATGC and λ-DNA, respectively. Bottom: DOS for λ-DNA using the same parameters as in the top panel.

Upon further increasing the disorder to $W = 2$, as shown in Fig. 7, the localisation lengths continue to decrease. Note that we observe a slight broadening of the bands and states begin to shift into the gap. We also see that the behaviour of random-ATGC and λ-DNA is quite similar and at these disorder strengths, even telomeric DNA follows the same trends. At $W = 5$, the localisation lengths have been reduced to a few base-pair separation distances and the differences between all 4 sequences are very small. The gap has been nearly completely filled as shown by the DOS, albeit with states which have a very small localisation length. This will become important later.

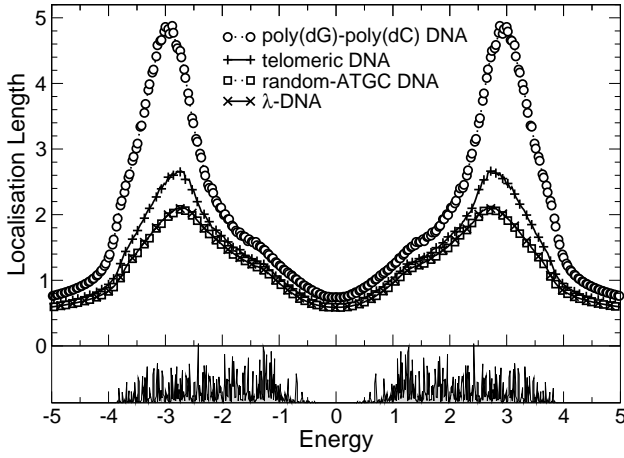Thus, in summary, we have seen that adding uniform disorder onto the backbone leads to a reduction of the

FIG. 8: Top: $\xi(E)$ as in Fig. 6 but with $W = 5$. Only every 2nd and 5th symbol is shown for random-ATGC and λ-DNA, respectively. Bottom: DOS for λ-DNA using the same parameters as in the top panel.

FIG. 9: Top: Energy dependence of the localisation lengths, $\xi(E)$, for poly(dG)-poly(dC), telomeric, random-ATGC and λ-DNA in the presence of *binary* backbone disorder with $W = 1$. Only every 2nd and 5th symbol is shown for random-ATGC and λ-DNA, respectively. Bottom: DOS for λ-DNA using the same parameters as in the top panel.

FIG. 10: Top: $\xi(E)$ as in Fig. 9 but with $W = 2$. Only every 2nd and 5th symbol is shown for random-ATGC and λ-DNA, respectively. Bottom: DOS for λ-DNA using the same parameters as in the top panel.

localisation lengths and consequently a reduction of the electron conductance. Strictly speaking, all 4 strands are insulators. However, their localisation lengths can remain quite large, larger than in many of the experiments. Thus even the localised electron can contribute towards a finite conductivity for these short sequences. In agreement with experiments, poly(dG)-poly(dC) DNA is the most prominent candidate.

B. DNA in an ionic solution

When in solution, the negatively charged oxygen on the backbone will attract cations such as Na^+. This will give rise to a dramatic change in local electronic properties at the oxygen-carrying backbone site, but not necessarily influence the neighbouring sites. The effects at each such site will be the same and thus in contrast to a uniform disorder used in section VI A, a binary distribution such as $\varepsilon_{i,q} = \pm W/2$ is more appropriate. For simplicity, we choose 50% of all backbone sites to be occupied $\varepsilon_{i,q} = -W/2$ while the other half remains empty with $\varepsilon_{i,q} = +W/2$. We note that a mixture of concentrations has been studied in the context of the Anderson model in Ref. [42].

In Fig. 9, we show the results for moderate binary disorder. In comparison with the uniformly disordered case of Fig. 6, we see that the localisation lengths have decreased further. This is expected because binary disorder is known to be very strong [42]. Also, the gap has already started to fill.

Increasing the disorder leads again to a decrease of ξ in the energy regions corresponding to the bands. Directly at $E = \pm W/2$, we observe 2 strong peaks in the DOS which is accompanied by reduced localization lengths. This peak corresponds to the infinite potential barrier or well at $E = -W/2$ or $+W/2$, respectively, as indicated
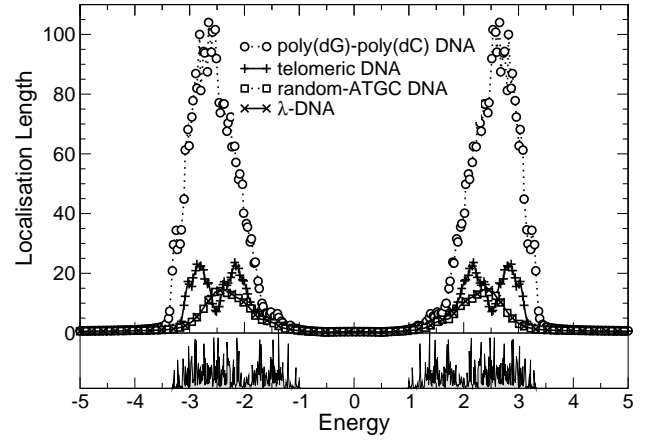
by Eq. (4). In Fig. 9, these peaks were not yet visible. We also see in Fig. 10 that the localisation lengths for states in the band centre start to increase to values $\gtrsim 1$. This trend continues for larger W as shown in Fig. 11. We see a crossover into a regime where the two original, weak-disorder bands have nearly vanished and states in the centre at $E = 0$ are starting to show an increasing localisation length *upon increasing the binary disorder*. A further increase in W eventually leads to the complete destruction of the original bands and the formation of a single band symmetric around $E = 0$ at about $W \sim 2.5$.
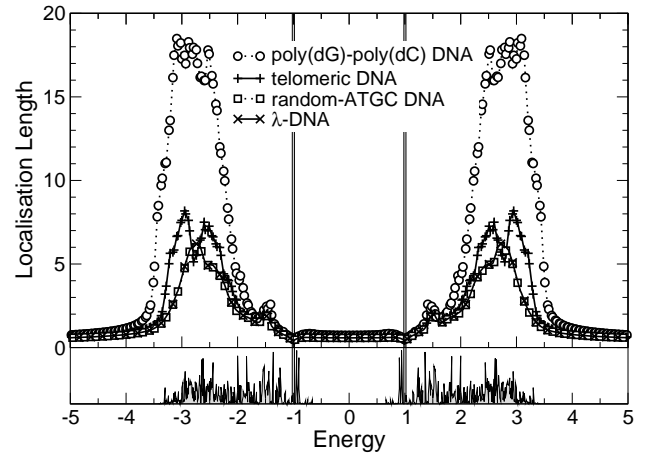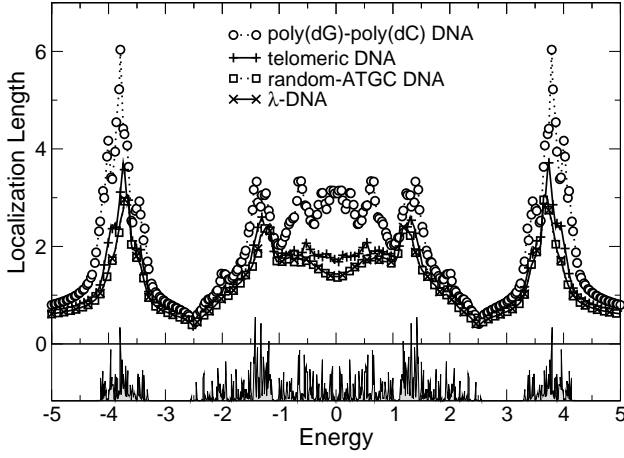
FIG. 11: Top: $\xi(E)$ as in Fig. 9 but with $W = 5$. Only every 2nd and 5th symbol is shown for random-ATGC and λ-DNA, respectively. Bottom: DOS for λ-DNA using the same parameters as in the top panel.
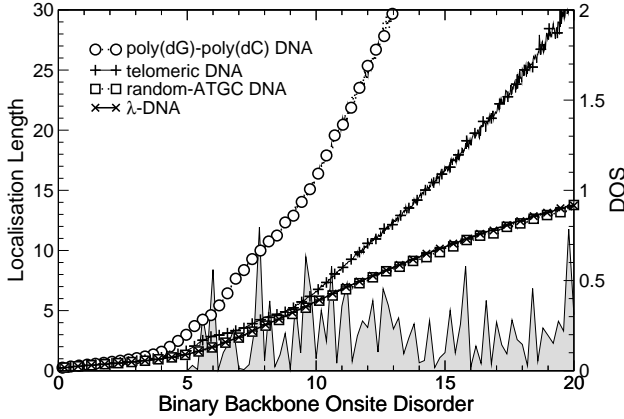

FIG. 12: Disorder dependence of ξ for poly(dG)-poly(dC), telomeric, random-ATGC and λ-DNA at $E = 0$. Only every 10th symbol is shown for all sequences. The shaded curve is the corresponding unnormalized DOS for λ-DNA.

C. Delocalisation due to disorder

The results of the previous section suggest that increasing the disorder in different regions of the energy will lead to different transport behaviour. Of particular interest is the region at $E = 0$. In Fig. 12 the variation of ξ as a function of binary disorder strength for all different sequences is shown. While $\xi < 1$ for small disorder, we see that upon increasing the disorder, states begin to appear and their localisation lengths increase for all DNA sequences. Thus we indeed observe a counter-intuitive *delocalisation* by disorder at $E = 0$. As before, poly(dG)-poly(dC) and telomeric disorder show the largest localisation lengths, whereas random-ATGC and λ-DNA give rise to a smaller and nearly identical effect. In Fig. 13 we show that this effect does not exist at $E = 3$, i.e. for energies corresponding to the formerly largest localisation
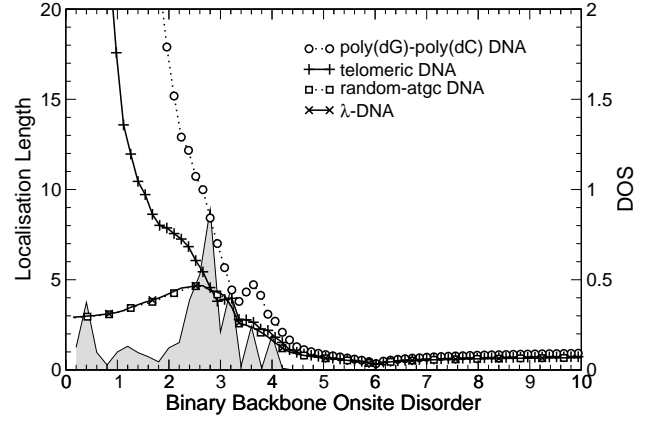

FIG. 13: $\xi(W)$ as in Fig. 12 but with $E = 3$. Only every 10th symbol is shown for all DNA sequences. The shaded curve is the corresponding unnormalized DOS for λ-DNA.
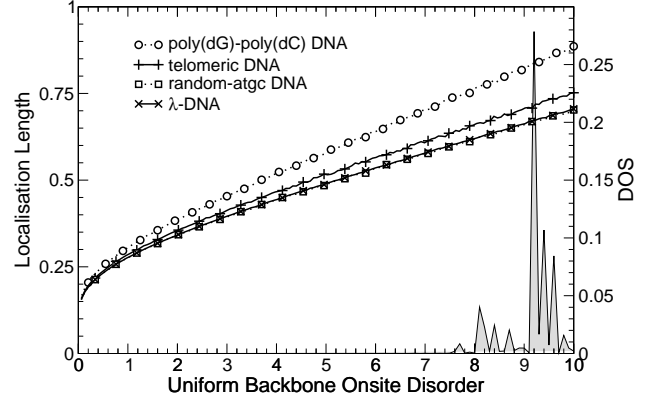

FIG. 14: $\xi(W)$ as in Fig. 12 but with uniform disorder at $E = 0$ and for the *fishbone model*. Only every 10th symbol is shown for all DNA sequences. The shaded curve is the corresponding unnormalized DOS for λ-DNA.

lengths. Rather, at $E = 3$, the localisation lengths for all DNA sequences quickly drop to $\xi \sim 1$. The delocalisation effect is also observed for uniform disorder, but is much smaller. As shown in Fig. 14, the enhancement is up to about $\xi = 1$ for the fishbone model (1). Results for the ladder model (2) are about 1.7 times larger.

This surprising delocalisation-by-disorder behaviour can be understood by considering the effects of disorder at the backbone for the effective Hamiltonians (3) and (4). At $E = 0$, the onsite potential correction term $(t_i^q)^2/(\varepsilon_i^q - E)$ will *decrease* upon increasing the ε_i^q values. For binary disorders $\varepsilon_i^q = \pm W/2$, this holds for $|\varepsilon_i^q| > |E|$ as shown in Fig. 13. However, for large $|E|$, the localisation lengths decrease quickly due to the much smaller density of states. Thus the net effect is an eventual decrease (or an only very small increase) of ξ for large E. Note the dip at $|\varepsilon_i^q| = E = 3$ in the figure, which corresponds to the effective $\varepsilon_i = \infty$, i.e. an infinitely strong trap yielding extremely strong localisation. For uniform disorder $\varepsilon_i^q \in [-W/2, W/2]$ — and generally any disor-

der with compact support around $E = 0$ — the above inequality is never full-filled and even for $E = 0$ we will find small $\varepsilon_i^q \sim 0$ such that we have strong trapping and localisation.

VII. INVESTIGATING THE LOCAL PROPERTIES OF THE SEQUENCES

A. Variation of ξ along the DNA strand

In the preceding sections, we had computed estimates of the localisation length ξ for complete DNA strands, i.e. the ξ values are *averages*. However, the biological function of DNA clearly depends on the local structure of the sequence in a paramount way. After all, only certain parts of DNA code for proteins, while others do not. In addition, the exact sequence of the bases specifies the protein that is to be assembled. Thus, in order to gain access to the local properties, we have performed computations of ξ on subsequences of complete DNA strands. We start by artificially restricting ourselves to finite windows of length $K = 10, 30, 50, 100, 200, 500, 1000$ and compute the localisation lengths $\xi_K(r)$ where $r = 1, 2, \ldots, L - K$ denotes the starting position of the window of length K.

In order to see how the exact sequence determines our results, we have also randomly permuted (scrambled) the λ-DNA sequence so that the content of A, T, G, and C bases is the same, but their order is randomised. Differences in the localisation properties should then indicate the importance of the exact order. From the biological information available on bacteriophage λ-DNA, we compute the localisation length for the coding regions [14] and then for window lengths K that correspond exactly to the length of each coding region. Again, if the electronic properties — as measured by the localisation length — are linked to biological content, we would expect to see characteristic differences.

In Figs. 15 and 16, we show results for $K = 100$ and 1000, respectively. From Fig. 15, we see from $P(\xi)$ that the localisation lengths for λ-DNA are mostly distributed around 15–20, but $P(\xi)$ has a rather long tail for large ξ. However, there are some windows where the localisation lengths exceed even the size of the window $K = 100$. Thus at specific positions in the DNA sequence, the system appears essentially extended with $\xi > K$. On the other hand, the distribution $P(\xi)$ is identical when instead of λ-DNA, we consider scrambled DNA. Therefore the presence of such regions is not unique to λ-DNA. The results from windows positioned at the coding part of λ-DNA appear statistically similar to the complete sequence, i.e. including also the non-coding regions. This suggests that with respect to the localisation properties there is no obvious difference between λ-DNA and scrambled λ-DNA as well as coding and non-coding regions. We emphasise that similar results have been obtained for a DNA sequence constructed from the SARS corona-viral data.
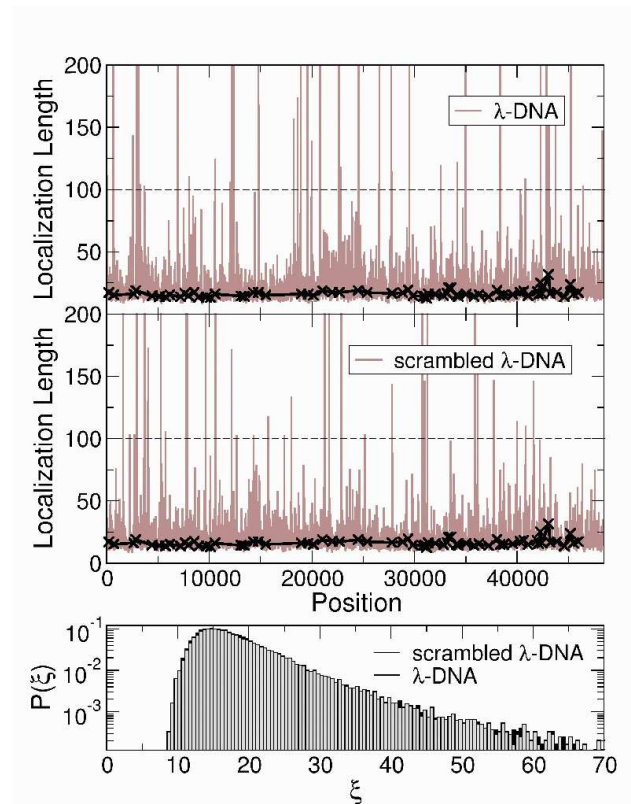

FIG. 15: Top: Variation of the localisation lengths for a sliding window of length $K = 100$ as a function of window starting position for λ-DNA at $E = 3$. The black crosses ($\times$) denote results for windows corresponding to the coding sequences of λ-DNA only. The dashed horizontal line denotes K. Middle: Same as in the top panel but with randomly scrambled λ-DNA. Bottom: Normalised distribution functions $P(\xi)$ for the localisation lengths ξ of λ- (black) and scrambled-λ-DNA (grey).

In Fig. 15, we repeat these calculations but with $K = 1000$. Clearly, $P(\xi)$ is peaked again around 15–20 and this time has no tail. In all cases, $K > \xi$. Again, the results for scrambled DNA are different in each window, and now even $P(\xi)$ is somewhat shifted with respect to λ-DNA.

Thus in conclusion, we do not see significant differences between λ-DNA and its scrambled counter part. Moreover, there appears to be no large difference between the localisation lengths measured in the coding and the non-coding sequences of bacteriophage λ-DNA. This indicates that the average ξ values computed in the previous sections is sufficient when considering the electronic localisation properties of the 4 complete DNA sequences.

B. Computing correlation functions

As shown in the last section, the spatial variation of ξ for a fixed window size is characteristic of the order
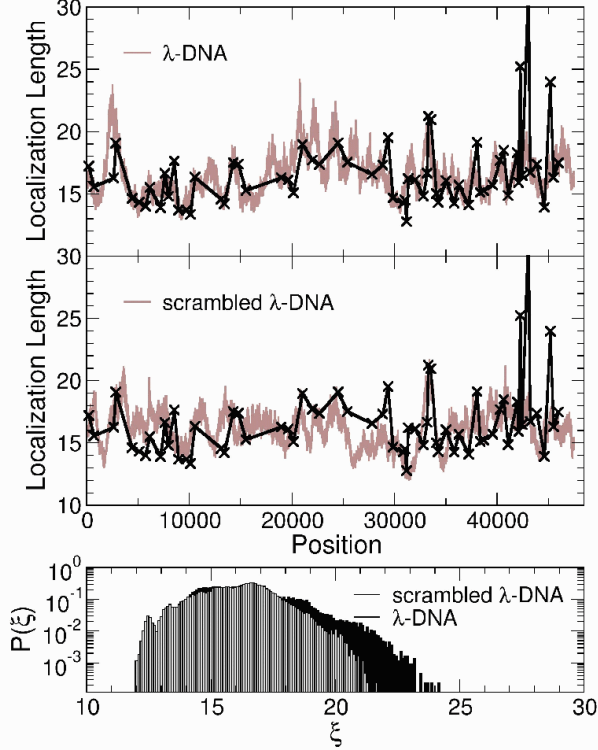
FIG. 16: Variation of the localisation lengths for a sliding window of length $K = 1000$ at $E = 3$ as in Fig. 15. Middle: Same as in the top panel but with randomly scrambled λ-DNA. Bottom: Normalised distribution functions $P(\xi)$ for the localisation lengths ξ of λ- (black) and scrambled-λ-DNA (grey).

of bases in the DNA sequence. Thus we can now study how this biological information is retained at the level of localisation lengths. In order to do so, we define the correlation function

$$\mathrm{Cor}(k) = \frac{\sum_{i=1}^{n-k} \left[\xi(r_i) - \langle\xi\rangle\right]\left[\xi(r_{i+k}) - \langle\xi\rangle\right]}{\sum_{i=1}^{n} \left[\xi(r_i) - \langle\xi\rangle\right]^2} \qquad (5)$$

where $\langle\xi\rangle = \sum_{i=1}^{n} \xi(r_i)/n$ is ξ averaged over all $n = L - (K-1)$ windows for each of which the individual localisation lengths are $\xi(r_i)$.

In Fig. 17 we show the results obtained for λ-DNA with windows of length 10, 100 and 1000. We first note that $\mathrm{Cor}(k)$ drops rapidly until the distance k exceeds the window width K (see the inset of Fig. 17). For $k > K$, $\mathrm{Cor}(k)$ fluctuates typically between ± 0.2 and there is a larger anti-correlation for base-pair separations of about $k = 8000$. We note that such large scale features are not present when considering scrambled λ-DNA instead.
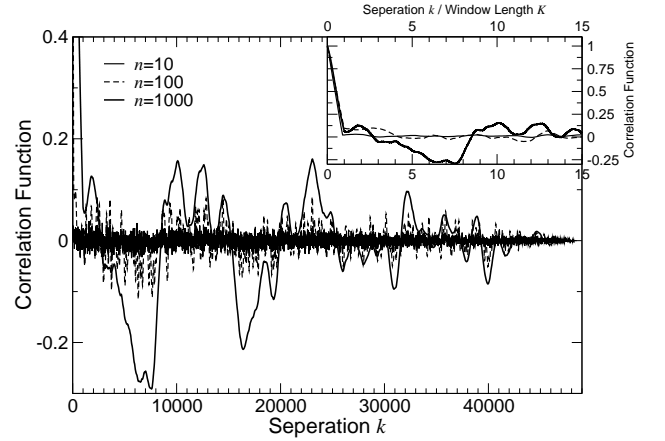

FIG. 17: $\mathrm{Cor}(k)$ as defined in Eq. (5) for λ-DNA and $K = 10$, 100, and 1000 at $E = 0$. The inset shows the same date but plotted as a function of normalized seperation k/K.

VIII. DISCUSSION

The fishbone and ladder models studied in the present paper give qualitatively similar results, i.e. a gap in the DOS on the order of the hopping energies to the backbone, extended states for periodic DNA sequences and localised states for any non-zero disorder strength. Thus at $T = 0$, our results suggest that DNA is an insulator unless perfectly ordered. Quantitatively, the localisation lengths ξ computed for the ladder model are larger than for the fishbone model. Since we are interested in these non-universal lengths, the ladder model is clearly the more appropriate model.

The localisation lengths measure the spatial extent of a conducting electron. Our results suggest — in agreement with all previous considerations — that poly(dG)-poly(dC) DNA allows the largest values of ξ. Even after adding a substantial amount of disorder, poly(dG)-poly(dC) DNA can still support localization lengths of a few hundred base-pair seperation lengths. With nanoscopic experiments currently probing at the most a few dozen bases, this suggests that poly(dG)-poly(dC) DNA will appear to be conducting in these experiments.

Furthermore, telomeric DNA is a very encouraging and interesting naturally occuring sequence because it gives very large localisation lengths in the weakly disordered regime. Nevertheless, we find that all investigated, non-periodic DNA sequences such as, e.g. random-ATGC and λ-DNA, give localised behaviour even in the clean state. This indicates that they are insulating at $T = 0$.

When the effects of the environment, modelled by their potential changes on the backbone, are included, we find that the localisation lengths in the two bands decrease quickly upon increasing the disorder. Nevertheless, depending on the value of the Fermi energy, the resulting ξ values can still be 10-20 base-pairs long. While this may not give metallic behavior, it can still result in a finite current for small sequences. We also note that these

distances are quite close to those obtained from electron-transfer studies.

The backbone disorder also leads to states moving into the gap. Therefore the environment prepared in the experiments determines the gap which is being measured. Furthermore, the localisation properties of the states in the former gap are drastically different from those in the 2 bands. Increasing the disorder leads to an increase in the localization lengths and thus potentially larger currents. This is most pronounced for binary disorder, taken to model the adhesion of cations in solution. Thus within the 2 models studied, we find that their transport properties are in a very crucial way determined by the environment. Differences in experimental set-up such as measurements in 2D surfaces or between elevated con-tacts are likely to lead to quite different results.

As far as the correlations within biological λ-DNA are concerned, we see only a negligible difference between the localisation properties of the coding and non-coding parts. However, this is clearly dependent on the chosen energy and the particular window lengths used. Investigations on other DNA sequences are in progress.

[1] SPARTAN version 5.0, User's Guide. Wavefunction Inc., 18401 Von Karman Ave., Suite 370 Irvine, CA 92612.

[2] Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J. Watson. 1994. Molecular Biology of the Cell. Garland, New York.

[3] Asabaeva, A., and M. Tang. 2000. Electrical conductivity in oriented DNA. *National Nanofabrication Users Network Newsletter.* :56–57.

[4] Barnett, R. N., C. L. Cleveland, A. Joy, U. Landman, and G. B. Schuster. 2001. Charge migration in DNA: ion-gated transport. *Science.* 294:567–571.

[5] Berlin, Y. A., A. L. Burin, and M. A. Ratner. 2000. On the long-range charge transfer in DNA. *The Journal of Physical Chemistry.* 104:443–445.

[6] Bhalla, V., R. P. Bajpai, and L. M. Bharadwaj. 2003. DNA electronics. *European Molecular Biology reports.* 4:442–445.

[7] Biswas, P., P. Cain, R. A. Römer, and M. Schreiber. 2000. Off-diagonal disorder in the Anderson model of localization. *phys. stat. sol. (b).* 218:205–209. ArXiv: cond-mat/0001315.

[8] Bixon, M., B. Giese, S. Wessely, T. Langenbacher, M. E. Michel-Beyerle, and J. Jortner. 1999. Long-range charge hopping in DNA. *PNAS.* 96:11713.

[9] Boon, E., A. Livingston, N. Chmiel, S. David, and J. Barton. 2003. DNA-mediated charge transport for DNA repair. *Proc. Nat. Acad. Sci.* 100:12543–12547.

[10] Braun, E., Y. Eichen, U. Sivan, and G. Ben-Yoseph. 1998. DNA-templated assembly and electrode attachment of a conducting silver wire. *Nature.* 391:775–778.

[11] Bruinsma, R., G. Gruner, M. R. D'Orsogna, and J. Rudnick. 2000. Fluctuation-facilitated charge migration along DNA. *Phys. Rev. Lett.* 85:4393–4396.

[12] Cuenda, S., and A. Sanchez. 2004. Disorder and fluctuations in nonlinear excitations in DNA. ArXiv: q-bio.BM/0403003v1.

[13] Cuniberti, G., L. Craco, D. Porath, and C. Dekker. 2002. Backbone-induced semiconducting behavior in short DNA wires. *Phys. Rev. B.* 65:241314–4.

[14] Daniels, D. L., J. L. Schroeder, W. Szybalski, F. Sanger, and F. R. Blattner. 1983. LAMBDA II:469-517, chapter Appendix I: A molecular map of colphase lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor.

[15] Davies, O. R., and J. E. Inglesfield. 2004. Embedding methods for conductance in DNA. *Phys. Rev. B.* 69:195110–13.

[16] Dekker, C., and M. A. Ratner. 2001. Electronic properties of DNA. *Physics World.* 14:29.

[17] Delaney, S., and J. K. Barton. 2003. Long-range DNA charge transport. *J. Org. Chem.* 68:6475–6483.

[18] Dyson, F. J. 1953. The dynamics of a disordered linear chain. *Phys. Rev.* 92:1331–1338.

[19] Eilmes, A., R. A. Römer, and M. Schreiber. 1998. The two-dimensional Anderson model of localization with random hopping. *Eur. Phys. J. B.* 1:29–38.

[20] Endres, R. G., D. L. Cox, and R. P. Singh. 2004. The quest for high-conductance DNA. *Rev. Mod. Phys.* 76:195–214.

[21] Fink, H.-W., and S. C. 1999. Electrical conduction through DNA molecules. *Nature.* 398:407–410.

[22] Frahm, K., A. Müller-Groeling, J. L. Pichard, and D. Weinmann. 1995. Scaling in interaction-assisted coherent transport. *Europhys. Lett.* 31:169.

[23] Garzon, I. L., E. Artacho, M. R. Beltran, A. Garcia, J. Junquera, K. Michaelian, P. Ordejon, C. Rovira, D. Sanchez-Portal, and J. M. Soler. 2001. Hybrid DNA-gold nanostructured materials: an ab-initio approach. *Nanotechnology.* 12:126–131.

[24] Gutierrez, R., S. Mandal, and G. Cuniberti. 2004. Quantum transport in DNA wires: Influence of a strong dissipative environment. ArXiv: cond-mat/0410660.

[25] Kelley, S. O., and J. K. Barton. 1999. Electron transfer between bases in double helical DNA. *Science.* 283:375–381.

[26] Kramer, B., and A. MacKinnon. 1993. Localization: theory and experiment. *Rep. Prog. Phys.* 56:1469–1564.

[27] MacKinnon, A. 1980. The conductivity of the one-dimensional disordered anderson model: a new numerical method. *J. Phys.: Condens. Matter.* 13:L1031–L1034.

[28] MacKinnon, A. 1985. The calculation of transport properties and density of states of disordered solids. *Z. Phys. B.* 59:385–390.

[29] MacKinnon, A. 1994. Critical exponents for the metal-insulator transition. *J. Phys.: Condens. Matter.* 6:2511–2518.

[30] MacKinnon, A., and B. Kramer. 1983. The scaling the-

ory of electrons in disordered solids: additional numerical results. *Z. Phys. B.* 53:1–13.

[31] Murphy, C. J., M. A. Arkin, Y. Jenkins, N. D. Ghatlia, S. Bossman, N. J. Turro, and J. K. Barton. 1993. Long-range photoinduced electron transfer through a DNA helix. *Science.* 262:1025–1029.

[32] Nakao, H., M. Gad, S. Sugiyama, K. Otobe, and T. Ohtani. 2003. Transfer-printing of highly aligned DNA nanowires. *J. Am. Chem. Soc.* 125:7162–7163.

[33] Ndawana, M. L., R. A. Römer, and M. Schreiber. 2004. Effects of scale-free disorder on the Anderson metal-insulator transition. *Europhys. Lett.* 68:678–684.

[34] Okahata, Y., T. Kobayashi, K. Tanaka, and M. Shimomura. 1998. Anisotropic electric conductivity in an aligned DNA cast film. *J. Am. Chem. Soc.* 120:6165–6166.

[35] O'Neil, M. A., C. Dohno, and J. K. Barton. 2004. Direct chemical evidence for charge transfer between photoexcited 2-aminopurine and guanine in duplex DNA. *Journal of the American Chemical Society Communications.* 126:1316–1317.

[36] O'Neil, P., and E. M. Fielden. 1993. Primary free radical processes in DNA. *Adv. Radiat. Biol.* 17:53–120.

[37] Pablo, P. J., F. Moreno-Herrero, J. Colchero, J. Gomez Herrero, P. Hererro, P. Baro, A. M. an Ordejon, J. M. Soler, and E. Artacho. 2000. Absence of dc-conductivity in λ-DNA. *Phys. Rev. Lett.* 85:4992–4995.

[38] Peng, C. K., S. Buldyrev, A. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley. 1992. Long-range correlations in nucleotide sequences. *Nature.* 356:168–171.

[39] Peyrard, M. 2004. Nonlinear dynamics and statistical physics of DNA. *Nonlinearity.* 17:R1–R40.

[40] Pichard, J.-L., and G. Sarma. 1981. Finite-size scaling approach to Anderson localisation. *J. Phys. C.* 14:L127–L132.

[41] Pichard, J.-L., and G. Sarma. 1981. Finite-size scaling approach to Anderson localisation: II. quantitative analysis and new results. *J. Phys. C.* 14:L617–L625.

[42] Plyushchay, I., R. A. Römer, and M. Schreiber. 2003. The three-dimensional anderson model of localization with binary random potential. *Phys. Rev. B.* 68:064201–8.

[43] Porath, D., A. Bezryadin, S. Vries, and C. Dekker. 2000. Direct measurement of electrical transport through DNA molecules. *Nature.* 403:635–638.

[44] Porath, D., G. Cuniberti, and R. Di Felice. 2004. Charge transport in DNA-based devices. *Topics in Current Chemistry.* 237:183.

[45] Rakitin, A., P. Aich, C. Papadopoulos, Y. Kobzar, A. Vendeneev, J. Lee, and J. Xu. 2001. Metallic conduction through engineered DNA: DNA nanoelectric building blocks. *Phys. Rev. Lett.* 86:3670–3673.

[46] Retel, J., B. Hoebee, J. E. F. Braun, J. T. Lutgerink, E. Van der Akker, H. Wanamarta, H. Joenjie, and M. V. M. Lafleur. 1993. Mutational specificity of oxidative DNA damage. *Mutation Res.* 299:165.

[47] Roche, S. 2003. Sequence dependent DNA-mediated conduction. *Phys. Rev. Lett.* 91:108101–4.

[48] Roche, S., D. Bicout, E. Maciá, and E. Kats. 2003. Long range correlation in DNA: Scaling properties and charge transfer efficiency. *Phys. Rev. Lett.* 91:228101–4.

[49] Römer, R. A., and M. Schreiber. 1997. The enhancement of the localization length for two-interacting particles is vanishingly small in transfer-matrix calculations. *Phys. Rev. Lett.* 78:4890.

[50] Römer, R. A., and M. Schreiber. 2003. The Anderson Transition and its Ramifications — Localisation, Quantum Interference, and Interactions, chapter Numerical investigations of scaling at the Anderson transition. Springer, Berlin, 3–19.

[51] Römer, R. A., C. Villagonzalo, and A. MacKinnon. 2002. Thermoelectric properties of disordered systems. *J. Phys. Soc. Japan.* 72:167–168. Suppl. A.

[52] Treadway, C. R., M. G. Hill, and J. K. Barton. 2002. Charge transport through a molecular π-stack: Double helical DNA. *Chemical Physics.* 281:409–428.

[53] Walet, N. R., and W. J. Zakrzewski. 2004. A simple model of the charge transfer in DNA-like substances. ArXiv: cond-mat/0402059v1.

[54] Wan, C., T. Fiebig, O. Schiemann, J. K. Barton, and A. H. Zewail. 2000. Femtosecond direct observation of charge transfer between bases in DNA. *Proc. Natl. Acad. Sci.* 97:14052–14055.

[55] Wang, H., J. P. Lewis, and O. F. Sankey. 2004. Band-gap tunneling states in DNA. *Phys. Rev. Lett.* 93:016401.

[56] Yamada, H., E. B. Starikov, D. Hennig, and J. F. R. Archilla. 2004. Localization properties of electronic states in polaron model for poly(dg)-poly(dc) and poly(da)-poly(dt) DNA polymers. ArXiv: cond-mat/0407148v1.

[57] Yu, Z., and X. Song. 2001. Variable range hopping and electrical conductivity along the DNA double helix. *Phys. Rev. Lett.* 86:6018–6021.

[58] Zhang, W., and S. E. Ulloa. 2004. Extended states in disordered systems: Role of off-diagonal correlations. *Phys. Rev. B.* 69:153203.

[59] Zhang, W., and S. E. Ulloa. 2004. Structural and dynamical disorder and charge transport in DNA. *Microelectronics Journal.* 35:23–26.

[60] Zhong, J. 2003. Electronic transport in DNA molecules with backbone disorder. *Phys. Rev. B.* (private communication).

[61] We note that Ref. [53] assumes transport is via the sugar-phosphate backbone.

[62] The results for the fishbone and ladder models are qualitatively the same. Quantitatively, the ladder model results have a nearly twice larger localisation length. This factor approaches 2, if $t_{1,2} \to 0$. Therefore we will focus our discussion on the 2-channel ladder model.

[Bacteriophage lambda] Bacteriophage lambda, complete genome [gi|9626243|ref|NC_001416.1|[9626243]], Genbank Accession number NC_001416, http://www.ncbi.nlm.nih.gov/entrez/

[SARS] SARS coronavirus, complete genome [gi|30271926|ref|NC_004718.3|[30271926]], Genbank Accession number NC_004718, http://www.ncbi.nlm.nih.gov/entrez/

[CEN2] CEN2, Chromosome II centromere, http://www.yeastgenome.org/